

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 09161043 A

(43) Date of publication of application: 20.06.97

(51) Int. Cl

G06T 1/00
G06K 9/20
H04N 5/222

(21) Application number: 07322032

(71) Applicant: NEC CORP

(22) Date of filing: 11.12.95

(72) Inventor: TAKAHASHI KATSUHIKO
KAMIMURA TAKESHI

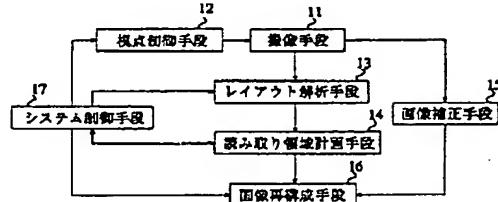
(54) DOCUMENT PICTURE INPUT DEVICE

(57) Abstract:

PROBLEM TO BE SOLVED: To efficiently acquire a document picture of an area larger than an image-pickup range by adding a control part to an image pickup device.

SOLUTION: A layout analyzing means 13 extracts the areas of a character, a chart and a photograph from a low resolution picture obtained by image- picking up the whole document, a reading area planning means 14 divides each area into ordered small areas and a picture reconstitution means 16 generates the picture of the whole document by sticking the pictures obtained by photographing each small area. As the picture reconstitution means 16 determines the position of a line within a character area from the result of layout analysis, correlative arithmetic for aligning is not necessitated concerning a direction orthogonal to a line direction which does not require the precision of aligning.

COPYRIGHT: (C)1997,JPO



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-161043

(43)公開日 平成9年(1997)6月20日

(51)Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
G 06 T 1/00			G 06 F 15/64	3 3 0
G 06 K 9/20	3 4 0		G 06 K 9/20	3 4 0 J
H 04 N 5/222			H 04 N 5/222	
			G 06 F 15/66	4 7 0 K

審査請求 有 請求項の数3 O L (全7頁)

(21)出願番号 特願平7-322032

(22)出願日 平成7年(1995)12月11日

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 高橋 勝彦

東京都港区芝五丁目7番1号 日本電気株式会社内

(72)発明者 上村 健

東京都港区芝五丁目7番1号 日本電気株式会社内

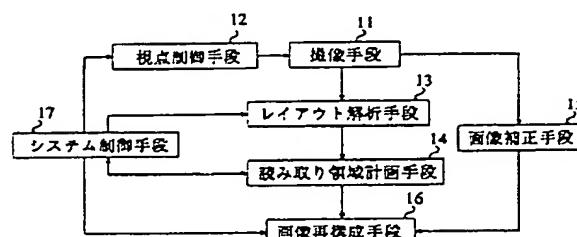
(74)代理人 弁理士 京本 直樹 (外2名)

(54)【発明の名称】 文書画像入力装置

(57)【要約】

【課題】 撮像装置に制御部を付加して、撮像範囲よりも広い領域の文書画像を効率よく獲得する。

【解決手段】 レイアウト解析手段13は文書全体を撮像した低解像度画像から文字、図表・写真領域を抽出し、読み取り領域計画手段14は各領域を順序づけされた小領域に分割し、画像再構成手段16は各小領域を撮影した画像を貼り合わせて文書全体の画像を生成する。画像再構成手段16は、文字領域内にある行の位置をレイアウト解析結果から決定するので、位置合わせの精度をそれほど要求されない行方向に直交する方向では位置合わせのための相関演算を行う必要がない。



1

【特許請求の範囲】

【請求項1】文書画像を取得する装置において、画像を取得するための撮像手段と、前記撮像手段の向き及びズームを制御する視点制御手段と、前記撮像手段が取得する文書の全体画像からレイアウトを解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果に基づき、文書中の文字領域及び図表・写真領域を複数の順番付けされた小領域に分割する読み取り領域計画手段と、前記撮像手段によって撮影した小領域画像の歪みを補正する画像補正手段と、前記画像補正手段によって補正された小領域画像を読み取り計画に基づき貼り合わせて、文書全体画像を生成する画像再構成手段と、前記撮像手段及び前記視点制御手段及び前記レイアウト解析手段及び前記画像再構成手段と通信してシステムを制御するシステム制御手段とから構成されることを特徴とする文書画像入力装置。

【請求項2】前記読み取り領域計画手段は、文字領域において、行方向に隣接した小領域が共通領域を持ち、行方向に直交する行並び方向に隣接する領域の境界が行間にくるように領域を設定し、前記画像再構成手段は、文字領域において、小領域画像の行方向における位置を行方向に隣接する小領域画像との相関から決定し、行並び方向における位置を前記読み取り領域計画手段を介して得られるレイアウト解析結果から決定することを特徴とする請求項1に記載の文書画像入力装置。

【請求項3】前記画像再構成手段により得られる画像から文字を認識して文字コードを生成する文字認識手段を備えることを特徴とする請求項1に記載の文書画像入力装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、スチルカメラやビデオカメラなどに視点制御機構を附加して実現した文書画像入力装置に関する。

【0002】

【従来の技術】装置の撮像範囲を超える領域の文書画像を獲得する方法として、特開平2-297148がある。特開平2-297148の発明では、互いに共通な領域を持つ複数の文書画像において相関を求める、相関が最大となる位置を画像貼り合わせのための概略位置とした後、前記文書画像中の本文文字列や見出し文字列等を照合して精密な貼り合わせを行い、全体画像を生成する。

【0003】また、文書画像のみではなく、一般の画像を貼り合わせて大規模画像を獲得する方法として特開平5-260264がある。特開平5-260264の発明では、2枚の画像において代表点の対応関係を求めてから画像にアフィン変換を施し、貼り合わせ処理を行って全体画像を構成する。

【0004】

2

【発明が解決しようとする課題】文書においては文章や図表といった領域間の境界となる背景領域（セパレータ）が存在し、そのセパレータで分割して入力を行うと精密な貼り合わせが不要となり効率的である。しかしながら、特開平2-297148や特開平5-260264に記載の発明は、文書全体をセパレータによって部分領域に分割して、各部分画像を獲得する機構を提供していない。ゆえに、これらの発明に基づいて大規模文書画像の入力装置を構成すると、必要以上に部分画像の撮影及び対応づけを行わなければならず、膨大な処理時間がかかるという問題点がある。

【0005】本発明は、以上のような問題点を解決するためになされたものであり、まずセパレータを検出して部分画像の撮影回数及び位置合わせのための画像間相関演算回数を少なくし、更に視点制御機構を備えることによって自動的に分割入力と貼り合わせを実行して、高精度文書画像を獲得する入力装置を提供することを目的とする。

【0006】

【課題を解決するための手段】上記の目的を達成するために、請求項1に記載の発明は、文書画像を取得する装置において、画像を取得するための撮像手段と、前記撮像手段の向き及びズームを制御する視点制御手段と、前記撮像手段が取得する文書の全体画像からレイアウトを解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果に基づき、文書中の文字領域及び図表・写真領域を複数の順番付けされた小領域に分割する読み取り領域計画手段と、前記撮像手段によって撮影した小領域画像の歪みを補正する画像補正手段と、前記画像補正手段によって補正された小領域画像を読み取り計画に基づき貼り合わせて、文書全体画像を生成する画像再構成手段と、前記撮像手段及び前記視点制御手段及び前記レイアウト解析手段及び前記画像再構成手段と通信してシステムを制御するシステム制御手段とから構成されることを特徴とする。

【0007】また、請求項2に記載の発明は、前記読み取り領域計画手段は、文字領域において、行方向に隣接した小領域が共通領域を持ち、行方向に直交する行並び方向に隣接する領域の境界が行間にくるように領域を設定し、前記画像再構成手段は、文字領域において、小領域画像の行方向における位置を行方向に隣接する小領域画像との相関から決定し、行並び方向における位置を前記読み取り領域計画手段を介して得られるレイアウト解析結果から決定することを特徴とする。

【0008】請求項3に記載の発明は、前記画像再構成手段により得られる画像から文字を認識して文字コードを生成する文字認識手段を備えることを特徴とする。

【0009】レイアウト解析手段は、文書の全体画像から、文字・図表写真領域や行の位置を抽出する。後段の処理はここで抽出された領域のみを処理するので、印刷

のない領域については画像の取り込み・貼り合わせを行う必要がない。読み取り領域計画手段は、抽出された各領域を小領域に分割する。但し、文字領域では、行方向に隣り合う小領域は重複領域を有し、且つ行並び方向においては行間に領域の境界がくるように領域を設定し、図表写真領域では隣接する領域が重複領域を有するよう網目状に領域を設定する。そして、撮像手段が撮影した各小領域画像を、画像再構成手段が貼り合わせて全体画像を生成する。このとき文字領域内では、各行の位置をレイアウト解析結果に基づいて決定し、行方向に隣接した小領域画像間においてのみ相関演算による正確な貼り合わせ処理を行う。ゆえに、行並び方向に隣接する画像間では相関演算が不要となり、高速に全体画像を構成することができるようになる。図表や写真領域内では、従来法を用いて貼り合わせを行うので、図表領域などが混在する文書画像でも高速に高精細の全体画像を生成することができる。

【0010】

【発明の実施の形態】以下、図面を参照して本発明の実施例を詳細に説明する。

【0011】本発明の実施例を図1に示す。本実施例は、CCDカメラから構成される撮像手段11と、CCDカメラを制御する視点制御手段12と、文書全体を撮影した画像からレイアウトを解析するレイアウト解析手

文字領域32 左上点(X1, Y0)

文字領域33 左上点(X0, Y2)

文字領域34 左上点(X0, Yn)

と各領域における行の位置：

行35 左上点(X1, Y0) 右下点(X3, Y1)

行36 左上点(X0, Y2) 右下点(X4, Y3)

.....

行37 左上点(X0, Y(z-1)) 右下点(X2, Yz)

が抽出される。

【0015】読み取り領域計画手段14の動作を図4を用いて説明する。本手段は、行グループ化手段41及び行グループ分割手段42及び読み取り順序決定手段43とから構成される。行グループ化手段41はレイアウト解析手段13の認識結果から、1つの小領域画像に取り

行35 左上点(X1*8, Y0*8) 右下点(X3*8, Y1*8)

行36 左上点(X0*8, Y2*8) 右下点(X4*8, Y3*8)

.....

行37 左上点(X0*8, Y(z-1)*8) 右下点(X2*8, Yz*8)

)

に位置すればよい。ここでさらに、レイアウト解析に用いた画像上の文字高を6ドット、行間隔を6ドットと仮定すると、作成する高精細画像上の文字高及び行間隔はそれぞれおよそ48ドットとなる。この時、小領域画像の行並び方向には5行程度を撮像可能であるが、文字欠けなく撮影できるのは最大4行である。そこで、このような場合には4行を1グループとして行グループを

段13と、レイアウト解析手段13の出力結果から文書中の文字領域を順序づけられた複数の小領域に分割する読み取り領域計画手段14と、撮像手段11より取得できる小領域画像の歪みを補正する画像補正手段15と、補正後の小領域画像を貼り合わせて文書全体の高精細画像を構成する画像再構成手段16と、システム全体を制御するシステム制御手段17から構成される。本実施例で想定するスタンド型文書画像入力装置の外観を図2に示す。

【0012】撮像手段11はズーム機能を持ったCCDカメラから構成される。本手段は、視点制御手段12により信号を受けてカメラのズームを制御できる。この機能を使って文書全体の画像を撮影したり、文書の一部を撮影することができる。

【0013】視点制御手段12は撮像手段11を固定する雲台であり、カメラをパン・チルトさせる機能、及び撮像手段11にズーム信号を送信しカメラのズームを制御する機能を有する。

【0014】レイアウト解析手段13の説明を図3を用いて行う。本手段は、撮像手段11がとらえた文書全体の画像31から、文書のレイアウトを解析する。レイアウト情報としては、文字・図表写真領域や行の位置を抽出する。図3の例では、3つの文字領域：

右下点(X3, Y1)

右下点(X4, Ym)

右下点(X4, Yz)

込む行数を決定し、行をグループ化する。ここで、グループ化の方法を説明する。簡単のために、CCDカメラの画素数を $640 * 400$ 、レイアウト解析時の画像の解像度を50DPI、再構成画像及び小領域画像を 400DPI の解像度で作成・撮影するものとする。この場合、図3に示した各文字行は高精細画像中では、

右下点(X3*8, Y1*8)

右下点(X4*8, Y3*8)

生成する。このようにしてグループ化した例を図5に示す。

【0016】また、図3に示した文書画像の文字高、行間隔がともに7ドットであったとすると、この場合は3行分の文字列が一枚の高精細画像に収録できる。但し、図6のように行並び方向に隣り合うグループの境界を完全に接触させてしまうと、完全に収まる行数が2行にな

る行グループ62が発生してしまう。文書画像において行間はあまり重要ではないので、このような場合には図7に示すように、常に行グループを行の上端から設定するようとする。図6、図7の場合では行グループの数に変化はないが、一般に行が多い文字領域であればこの操作によって行グループの数を減らすことができる。

【0017】行グループ分割手段42について説明する。本手段は、行グループ化手段41によってグループ化された領域を行方向に分割する。CCDのサイズが 640×400 、小領域画像の解像度が400DPIならば、各小領域の幅は $640 / 400$ inch以下に設定する。また、図8に示すように、隣り合う領域がある程度重複するように分割を行い、画像貼り合わせ時の相関計算に利用する。

【0018】読み取り順序決定手段43は行グループ分割手段42によって分割された小領域の読み取り順序を決定する。効率よく読み取るには、隣接する領域を連続して読むようにすればよい。効率よい読み取り順序の例として、図9のように原稿上部から下部に向かって蛇行させる方法などがある。

【0019】画像補正手段15は、撮像手段11より得られる小領域画像の歪みを補正する。画像の歪みは、主にレンズによる歪みと原稿に対するカメラの傾きから発生するので、これらの特性をあらかじめ測定しておき、これをを利用して画像を補正する。

【0020】画像再構成手段16について説明する。本手段は読み取り領域計画手段14で決定された小領域の画像を画像補正手段15より順番に取得し、これを貼り合わせて原稿全体の高精細画像を取得する。

【0021】まず、行グループにおいて読み取り順序がもっとも早い小領域画像が得られたならば、その位置、すなわち行の位置をレイアウト解析結果から決定する。その方法を図10を用いて説明する。レイアウト解析時の画像において、本領域に含まれる最上部行の左上端座標が(X_0, Y_0)であったとすると、高精細画像101では座標点104($N * X_0, N * Y_0$)に小領域画像102中の対応点103がくるようにすればよい。ここで、Nは高精細画像撮影時の解像度をレイアウト解析時の解像度で割った値を表す。本手法によれば、行方向及び行並び方向において、それぞれNドット前後のずれが高精細画像上で生じるが、その実寸値は1mm以下であるのでほとんど無視することができる。

【0022】次に、この画像と行方向に隣接する画像が得られたら、最初の画像とこれを重複領域の相関を利用して貼り合わせる。この方法を図11を用いて説明する。視点制御手段12がパン、チルト機構を備えていれば、小領域画像111を取得後カメラをパンするだけで行方向に並んだ小領域画像112を獲得することができる。2画像間の走査線はほぼ正しく対応していると考えられる。そのような場合には、行並び方向に黒画素

を計数し、2つのヒストグラム113及び114の形状がもつとも一致する位置を見つけることにより、2画像の貼り合わせ位置を決定することができる。また、カメラが微妙に上下し走査線の対応がずれる場合には、画像を2次元的にずらして最も画素値が一致する位置を抽出したり、文字切り出し処理やラベリング処理を施した結果の相関などを利用すればよい。このようにして、行方向に隣り合う画像を順次貼り合わせていき、行グループ画像115を完成させる。

【0023】行グループ画像が完成したら、次の行グループ画像について同様の処理を繰り返す。次の行グループ画像の位置もレイアウト解析結果、すなわち行グループの位置座標情報をもとに決定する。このように、本発明によれば、行の位置が微妙にずれるものの、行並び方向に隣接する領域間において位置合わせのための相関演算を行う必要はない。

【0024】システム制御手段17は、各処理部を起動したり、各処理部の処理結果を参照し次の行動を決定する機能を有する。

【0025】次に、処理の流れについて説明する。まず、システム制御手段17は視点制御手段12に文書全体の画像を撮像するための命令を送る。視点制御手段12はこれを受けてカメラの向きを変更し、ズームを変える指令を撮像手段11に送る。システム制御手段17は、撮像手段11が全体画像を撮影できるようになったらレイアウト解析手段13を起動する。レイアウト解析手段13は、撮像手段11より画像を取得し、レイアウト解析を行った後、終了信号をシステム制御手段17に送る。システム制御手段17は、次に読み取り領域計画手段14を起動する。読み取り領域計画手段14は文書画像全体を複数の小領域に分割して終了信号をシステム制御手段17に送る。システム制御手段17は読み取り領域計画手段14が計画した小領域の範囲パラメータを読み込み、各領域の画像が得られるように視点制御手段12にリクエストを送出する。視点制御手段12はこれにしたがってカメラの向き、ズームを変更する。システム制御手段17はカメラの位置決めが終了したら画像再構成手段16を起動する。画像再構成手段16は画像補正手段15が補正した小領域画像を貼り合わせて、文書全体の高精細画像を生成する。

【0026】本実施例では、横書き文書の例を示したが、縦書き文書でも同様に処理を進めることができる。さらに、縦書き・横書き混合の文書でも全く同様である。また、図表や写真が混在する原稿に対しては、それら領域を重複領域を有する小領域に網目状に分割した後、左上にある小領域の位置を文字領域の場合と同様にレイアウト解析結果から決定し、他の領域を従来の画像貼り合わせ手法によって貼り合わせればよい。

【0027】図1の構成に文字認識手段を追加した実施例を図12に示す。本実施例では、文字認識手段128

が画像再構成手段 126 のメモリに蓄積された高精細画像から文字を認識し、文字コードに変換する。文字認識手段 128 の起動はやはりシステム制御部 127 によってなされるが、そのタイミングは文書全体の画像を獲得した後でも、文字行グループの画像が得られる度でもよい。本発明によって構成される高精細画像は行並び方向に若干の誤差を含むが、文字コードを生成する場合には特にこの点を無視することができる。

【0028】

【発明の効果】本装置によれば文字や図表等の存在しない領域の高精度画像を撮像したり貼り合わせる必要がない。また、行の位置はレイアウト解析結果から決定するので、行並び方向に隣接する小領域画像間では相関演算をする必要がなく、従来の画像貼り合わせ装置に比べ画像再構成のための計算コストを大幅に削減することができる。

【0029】更に本装置によれば、紙面に印刷された文字ばかりでなく、屏風や立て看板などに書かれた文字のようにフラットベッドスキャナでは読みとることができるものでも自動入力することができる。

【図面の簡単な説明】

【図1】本実施例の構成を示すブロック図である。

【図2】文書画像入力装置の外観図の一例である。

【図3】レイアウト解析結果の一例である。

【図4】読み取り領域計画手段の内部構成を示すブロック図である。

【図5】行グループ化手段の処理結果の一例を示す図である。

【図6】好ましくない行グループ化手段の処理結果の一例を示す図である。

【図7】行グループ化手段の処理結果の一例を示す図である。

【図8】行グループ分割手段によって設定された小領域の一例を示す図である。

【図9】読み取り順序決定手段の処理結果を示す一例を示す図である。

【図10】小領域画像の高精細画像へのめ込み方を示す図である。

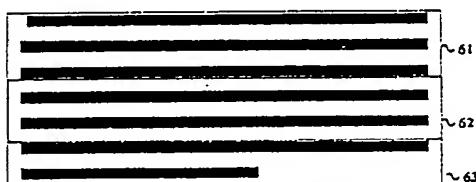
【図11】行並び方向に隣接する小領域画像の貼り合わせ方法を示す図である。

【図12】請求項3の実施例を示す図である。

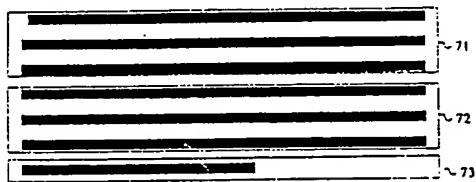
【符号の説明】

- 1 1 撮像手段
- 1 2 視点制御手段
- 1 3 レイアウト解析手段
- 1 4 読み取り領域計画手段
- 1 5 画像補正手段
- 1 6 画像再構成手段
- 1 7 システム制御手段
- 3 1 原稿
- 3 2 ~ 3 4 文字領域
- 3 5 ~ 3 7 行
- 4 1 行グループ化手段
- 4 2 行グループ分割手段
- 4 3 読み取り順序決定手段
- 4 4 読み取り領域計画手段
- 5 1 ~ 5 5 行グループ
- 6 1 ~ 6 3 行グループ
- 7 1 ~ 7 3 行グループ
- 8 1 ~ 8 5 小領域
- 8 6 行グループ
- 1 0 1 高精細画像
- 1 0 2 小領域画像
- 1 0 3 小領域画像上で最も上にある行の左上点
- 1 0 4 高精細画像上の点
- 1 1 1 小領域画像
- 1 1 2 1 1 1 の右側に隣接する小領域の小領域画像
- 1 1 3 小領域画像 1 1 1 の画素投影結果（ヒストグラム）
- 1 1 4 小領域画像 1 1 2 の画素投影結果（ヒストグラム）
- 1 1 5 行グループ画像
- 1 2 1 撮像手段
- 1 2 2 視点制御手段
- 1 2 3 レイアウト解析手段
- 1 2 4 読み取り領域計画手段
- 1 2 5 画像補正手段
- 1 2 6 画像再構成手段
- 1 2 7 システム制御手段
- 4 0 1 2 8 文字認識手段

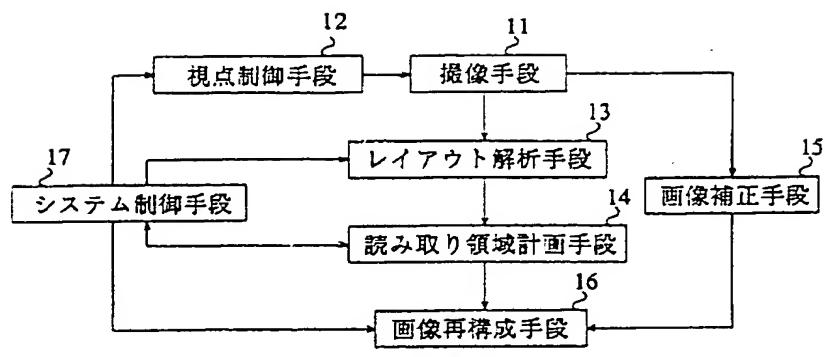
【図6】



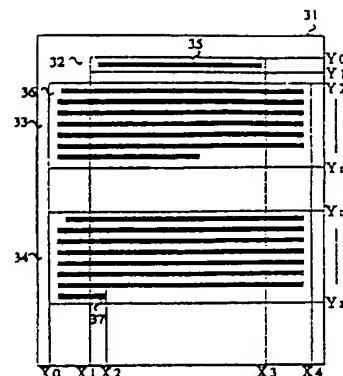
【図7】



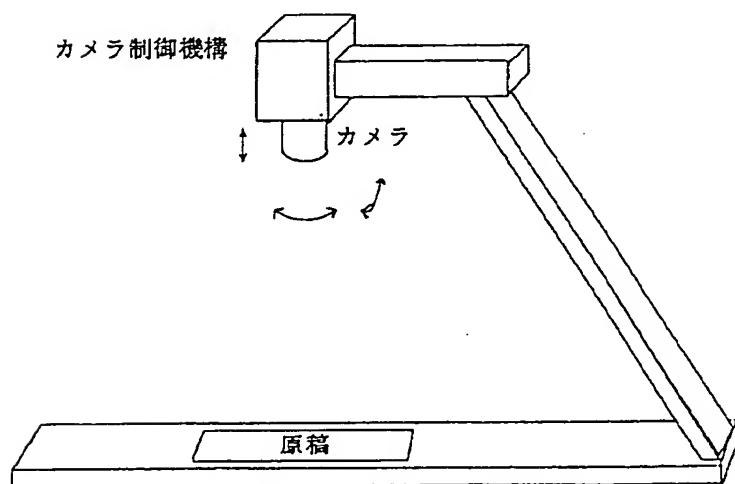
【図1】



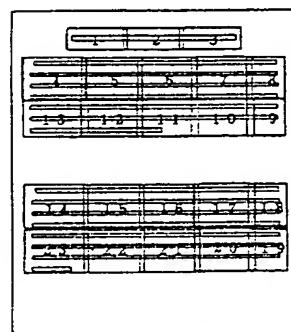
【図3】



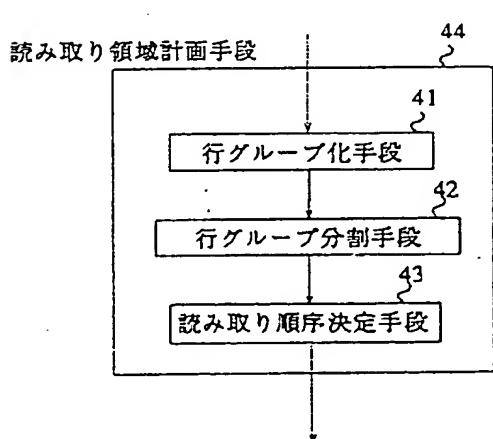
【図2】



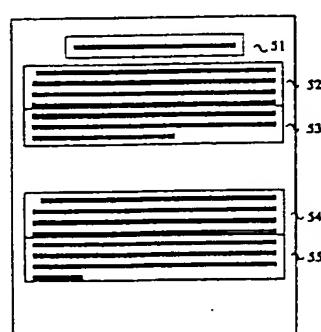
【図9】



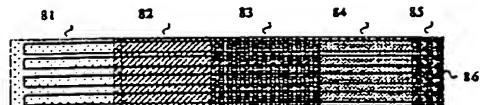
【図4】



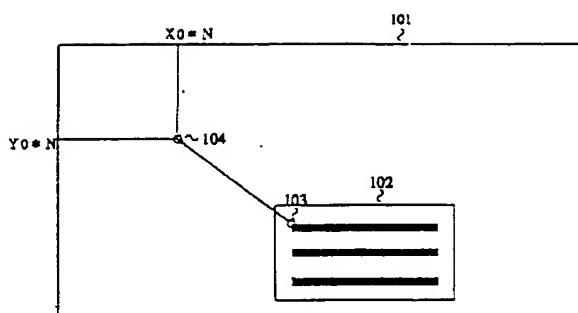
【図5】



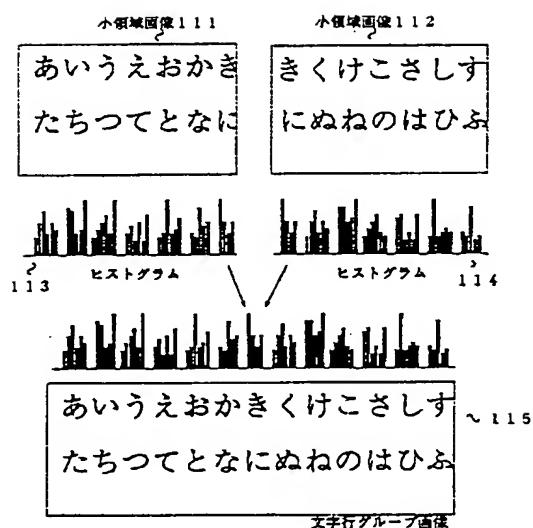
【図8】



【図10】



【図11】



【図12】

